JOURNAL OF INTERNATIONAL CRISIS AND RISK COMMUNICATION RESEARCH
ISSN: 2576-0017
2024, VOL 7, NO S7

DATA ANALYTICS MODELS FOR CRIME PREDICTION: A SUMMARY STUDY.

[1]William R. Insignares, [2]Emeldo Caballero B., [3]Pablo Carreño, [4]Pedro Jessid Pacheco Torres, [5]Randy Osorio.

1. Facultad de Ingeniería, Universidad Libre, carrera 46 48-170, wwwwww@unilibre.edu.co.
2. Facultad de Ingeniería, Universidad Libre, carrera 46 48-170, emeldo.caballerob@unilibre.edu.co.
3. Facultad de Ingeniería, Universidad Libre, Sede Bogota, pabloe.carrenoh@unilibre.edu.co
4. Facultad de Ingeniería, UAN, ppacheco606@uan.edu.co
5. SENA

SUMMARY
Machine-learning and deep-learning models are powerful tools for crime prediction. Using historical data, these models can identify patterns and trends, as well as estimate the likelihood of crime occurring in specific locations. Autonomous learning models include logistic regression, vector support machines (SVM), and decision trees, which are relatively simple and interpretable, but may have limitations with complex, nonlinear data. On the other hand, deep learning models, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversaries (GANs), can handle more complex data and capture intricate patterns, but they require large amounts of data and can be difficult to interpret. The choice of the right model will depend on the characteristics of the problem and the available data, taking advantage of the strengths of each approach.
Keywords: Machine Learning, Deep Learning, Crime Prediction, Neural Networks, Urban Security.

1. Introduction

Crime prediction models, such as those using machine learning techniques, offer advantages in improving preventive policing strategies and aiding resource allocation [1] [2]. These models can provide reasonable estimates of crime hotspots, which contributes to decision-making by law enforcement law enforcement agencies [3]. However, there are limitations, including difficulties in accurately predicting daily patterns of spatio-temporal crime at the urban district scale [4]. In addition, predictive models may come under criticism for their limited understanding of the social complexities of crime and potential biases in law enforcement [5]. Despite these drawbacks, predictive autonomous learning models hold promise for identifying individuals at high risk of criminal activity, contributing to the development of personalized rehabilitation strategies in psychiatry. More research is needed to address these limitations and biases, ensuring the effective and ethical use of crime prediction models for the benefit of society.

This work aims to present some of the most widely used data analytics models in the early detection of crimes, their advantages, disadvantages and limitations.

2. Methodology

To carry out this literature review, articles were searched in Science Direct databases. Keywords related to the topic of study were used and criteria were established to include only articles published between 2020 and 2023, in addition to some from previous years, prioritizing recognized scientific journals.

15 articles were chosen from a total of 25, the methodology, results and conclusions of each selected article were analyzed. The organization of the information was carried out in a synthesis matrix, identifying patterns and differences in the different models. The analysis focused on trends and areas that still need further research, for future studies in the field.

This systematic approach ensures that the review is relevant and current, thus contributing to knowledge on the topic.

3. Data Analytics Methods

To predict crime in cities and countries, machine learning and deep learning models have been used. These models use historical crime data to predict the likelihood of a crime occurring in a specific location, identifying patterns and trends in the crime data provided. Crime prediction helps law enforcement take preventative measures and reduce the crime rate in a city or country. Below are some machine learning and deep learning models used in crime prediction, along with their advantages and disadvantages.

3.1 Autonomous learning model.

These models can be classified according to the degree of human supervision required by the algorithm as supervised learning, unsupervised learning and reinforcement learning.

3.1.1. Supervised learning

This model uses a pre-labeled dataset for training and finding the desired solution [7, p.8] [8]. In this category, various activities are developed, highlighting the task of classification, where the models must generate a function that categorizes an input set, and the assignment of regression, which, although similar to classification, differs in that its output is continuous instead of discrete [6].

3.1.2. Unsupervised learning

In this category, the model does not have labeled data for training. In this algorithm, only the input data is known and the goal is to organize the data in such a way as to simplify its analysis. The most commonly used algorithms are Clustering, Principal Component Analysis and Anomaly Detection [7, p.10].

3.1.3. Reinforcement learning

This model is also known as Reinforcement. In this case, the algorithm interacts with a dynamic environment until it reaches a specific goal, accepting rewards or punishments as the problem evolves, by trying to autonomously formulate the best strategy for greater performance over time [7, p.14].

3.1.4. Other machine learning algorithms

There are various algorithms that vary in structure, input and output data, and computational complexity; among the most widely used Machine Learning models in crime prediction, especially for classification, are Logistic Regression, Vector Support Machines, or SVM, and Decision Trees.

3.1.4.1. Logistic regression

This type of regression is used to estimate the probability that a variable belongs to a specific class. If the probability is greater than 50%, the model predicts that the variable belongs to the positive class, otherwise, the prediction indicates that the class is negative [7, p.144]. This probability can be calculated with the equation:

$$\hat{p} = \sigma(X^T\theta) \qquad y = \begin{cases} 0 \ \ if\hat{p} < \ 0{,}5 \\ 1 \ \ if\hat{p} \leq \ 0{,}5 \end{cases}$$

where, is the estimated probability, σ the sigmoid function, the matrix of characteristics, θ the vector of parameters of the model, corresponds to the prediction. The model can be easily compressed and handle missing data. However, it can be prone to overfitting and is not suitable for non-linear data.$\hat{p}Xy$

3.1.4.2. Vector Support Machine

The Vector Support Machine model is part of the supervised learning methods for classification, regression or detection of endpoints. In these models, base functions centered on the training data points are defined and then a subset of these points is selected, which are called support vectors [6]. SVM predictions are particularly suitable for the classification of complex, but small or medium-sized datasets, and can perform linear and nonlinear classification tasks [9][7, p.155]. An important property of this type of model is that the parameters are determined as the solution to an optimization problem with a convex cost function, so that, despite the fact that a nonlinear problem is involved, the solution is relatively simple [6,8]. In short, this model is efficient in terms of memory and can handle non-linear data. However, it can be difficult to interpret and can be sensitive to parameter selection.

3.1.4.3. Decision trees

They are conceptually simple but powerful methods. The main idea is to select the prediction model based on the input variables. To obtain the structure of the tree, the input space is first divided into a set of rectangles, and then each region is assigned a simple model as a constant. This partition is then assigned its equivalent tree structure where each internal node denotes a test on one or more attributes, each branch represents a test output and the leaf nodes represent the classes [6][7, p.177]. In this way, a model combination method is obtained, where a single model is responsible for making predictions for a given point in the input space, and therefore can be described as a sequence of binary selections corresponding to the cross-sectional structure of the tree [6]. This model is easy to understand and visualize, and can handle missing data. However, it can be prone to overfitting and is not suitable for complex data.

3.2. Deep Learning Models

These models comprise convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative neural networks (GANs).

The Convolutional Neural Networks (CNN) model is efficient in terms of memory, is used for image processing and has been used for crime prediction. Its main advantage is its ability to identify patterns and features in images [10]. However, they require a lot of data and training time. Recurrent neural networks (RNNs) are used

for processing data streams, such as text, audio, or time series [11][12]. Their main advantage is their ability to model time dependence on data, making them suitable for sequential data, such as natural language. However, they can be prone to overfitting and require large amounts of data to train. RNNs have been used for crime prediction based on social media data. Another model is Generative Neural Networks (GANs), which can generate realistic synthetic data and can be used for image and video generation. However, it can be difficult to train and may require large amounts of data.

4. Models with generative AI

Data analytics models that employ generative artificial intelligence to predict crime are gaining attention as they improve in accuracy and efficiency. These models use historical crime data and advanced techniques, such as generative adversarial networks (GANs) and deep learning. Integrating human mobility patterns, from social media data and location services, has been shown to increase the effectiveness of predictions by 2% to 7% [13]. GANs, on the other hand, help balance classes in crime datasets by generating synthetic examples, improving performance on classification tasks [14]. However, the use of AI in crime prediction raises important ethical questions about transparency and the presumption of innocence, underscoring the need for safeguards to protect fundamental rights [15]. While generative AI models hold promise for increasing the accuracy of crime prediction, it is essential to address their ethical implications to ensure fair and accountable policing practices.

Analysis and Conclusion

Machine learning and deep learning models are widely used for crime prediction. These models use historical crime data to identify patterns and trends, and predict the likelihood of a crime occurring in a specific location.

Among the models of autonomous learning, logistic regression, vector support machines (SVM) and decision trees stand out. These models have the advantage of being easy to understand and interpret, and can handle missing data. However, they can be prone to overfitting, are not suitable for complex, non-linear data (in the case of logistic regression and decision trees), and can be sensitive to parameter selection (in the case of SVMs).

On the other hand, deep learning models, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial neural networks (GANs), have the ability to identify complex patterns and features in images, data streams, and structured data. Specifically, CNNs are efficient in terms of memory and can identify patterns in images, while RNNs are suitable for modeling temporal dependencies in sequential data, such as text and time series.

However, deep learning models also have limitations. These models can require large amounts of data and training time, can be prone to overfitting, and can be difficult to interpret and explain. In addition, GANs can be complicated to train and also require large amounts of data.

In summary, both machine learning and deep learning models offer advantages and disadvantages in predicting crime. Machine-learning models are simpler and more interpretable, but they may not be suitable for complex, non-linear data. On the other hand, deep learning models have the ability to identify complex patterns and features, but they can require large amounts of data and be difficult to interpret.

The choice of the most appropriate model will depend on the type of data available, the complexity of the problem, and the specific requirements of the crime prediction task. In some cases, it may be beneficial to combine different models or use ensemble learning approaches to build on the strengths of each model and mitigate its limitations.

References

[1] Jenga, K., Catal, C., & Kar, G. (2023). Machine learning in crime prediction. *Journal of Ambient Intelligence and Humanized Computing*, *14*(3), 2887-2913.

[2] Boqué, P., Saez, M., & Serra, L. (2022). Need to go further: using INLA to discover limits and chances of burglaries' spatiotemporal prediction in heterogeneous environments. *Crime Science*, *11*(1), 7.

[3] Hou, M., Hu, X., Cai, J., Han, X., & Yuan, S. (2022). An integrated graph model for spatial–temporal urban crime prediction based on attention mechanism. *ISPRS International Journal of Geo-Information*, *11*(5), 294.

[4] Rotaru, V., Huang, Y., Li, T., Evans, J., & Chattopadhyay, I. (2022). Event-level prediction of urban crime reveals a signature of enforcement bias in US cities. *Nature human behaviour*, *6*(8), 1056-1068.

[5] Watts, D., de Azevedo Cardoso, T., Librenza-Garcia, D., Ballester, P., Passos, I. C., Kessler, F. H., ... & Kapczinski, F. (2022). Predicting criminal and violent outcomes in psychiatry: a meta-analysis of diagnostic accuracy. Translational psychiatry, 12(1), 470.

[6] C. M. Bishop, Pattern Recognition and Machine Learning. NewYork: SpringerScience+BusinessMedia,2006.

[7] Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.".

[8] Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

[9] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, *2*(2), 121-167.
[10] Teuwen, J., & Moriakov, N. (2020). Convolutional neural networks. In *Handbook of medical image computing and computer assisted intervention* (pp. 481-501). Academic Press.
[11] Mandic, D. P., & Chambers, J. (2001). *Recurrent neural networks for prediction: learning algorithms, architectures and stability*. John Wiley & Sons, Inc..
[12] Chan, K. Y., Abu-Salih, B., Qaddoura, R., Ala'M, A. Z., Palade, V., Pham, D. S., ... & Muhammad, K. (2023). Deep Neural Networks in the Cloud: Review, Applications, Challenges and Research Directions. Neurocomputing, 126327.
[13] Wu, J., Abrar, S. M., Awasthi, N., Frias-Martinez, E., & Frias-Martinez, V. (2022). Enhancing short-term crime prediction with human mobility flows and deep learning architectures. EPJ data science, 11(1), 53.
[14] Engelmann, J., & Lessmann, S. (2021). Conditional Wasserstein GAN-based oversampling of tabular data for imbalanced learning. Expert Systems with Applications, 174, 114582.
[15] Sachoulidou, A. (2023). Going beyond the "common suspects": to be presumed innocent in the era of algorithms, big data and artificial intelligence. Artificial Intelligence and Law, 1-54.